DOI:
10.1002/eat.22773

Document Version
Peer reviewed version

Title

Sudden Gains in the Outpatient Treatment of Anorexia Nervosa: A Process-Outcome Study

Authors

Miss Anna Cartwright, MSc[1]

Miss Yat Ping Cheng, MSc[1]

Professor Ulrike Schmidt, MD PhD FRCPsych[2*]

Professor Sabine Landau, PhD[3*]

Affiliations

[1]: King's College London, Institute of Psychiatry, Psychology and Neuroscience, Department of Psychosis Studies, London, UK

[2]: King's College London, Institute of Psychiatry, Psychology and Neuroscience, Department of Psychological Medicine, London, UK

[3]: King's College London, Institute of Psychiatry, Psychology and Neuroscience, Department of Biostatistics and Health Informatics, London, UK

* Joint last authors

Acknowledgements

This paper represents independent research part-funded by the National Institute for Health Research (NIHR) Biomedical Research Centre at South London and Maudsley NHS Foundation Trust and King's College London. The views expressed are those of the authors and not necessarily those of the NHS, the NIHR or the Department of Health.

Word count for abstract: 228 words

Word count for manuscript: 4090 words

SUDDEN GAINS IN ANOREXIA NERVOSA

Abstract

Objective: Sudden gains (SGs), broadly defined as sudden symptom reductions occurring between two consecutive treatment sessions, have been associated with improved treatment outcomes in anxiety and depression. The present study is the first to formally define SGs in anorexia nervosa and explore the characteristics, demographic and baseline clinical predictors, and clinical impact of SGs in anorexia nervosa. **Method:** This is a secondary analysis of data from 89 outpatients with broadly defined anorexia nervosa who received one of two psychotherapeutic interventions as part of the MOSAIC trial (Schmidt et al., 2015). SGs were defined using session-by-session Body Mass Index (BMI) measures. This study investigated whether SGs were associated with changes in BMI, eating disorder symptomology, general psychopathology, and psychosocial impairment between baseline and 6, 12 and 24 months follow-up. **Results:** SGs, experienced by 61.8% of patients, mostly occurred during the early and middle phases of treatment. A larger proportion of SGs predicted larger increases in BMI between baseline and 6, 12 and 24 months follow-up. Amongst those experiencing at least one SG, fewer days between baseline and a patient's first SG predicted a larger increase in BMI between baseline and both 6 and 12 months follow-up. The proportion and timing of SGs did not predict changes in other outcome measures. **Discussion:** SGs in BMI during the outpatient treatment of anorexia nervosa are clinically useful predictors of longer-term weight outcomes.

Key Words: Feeding and Eating Disorders Anorexia Nervosa, Treatment Outcome, Body Mass Index, Weight Gain,

Sudden Gains in the Outpatient Treatment of Anorexia Nervosa: A Process-Outcome Study

Anorexia Nervosa (AN) is associated with high disability and mortality, posing major psychological and economic burden (Gatt et al., 2014; Stuhldreher et al., 2012). International treatment guidelines recommend psychotherapy for adults with AN (National Collaborating Centre for Mental Health., 2004) and high quality large scale randomized controlled trials (RCTs) are now beginning to emerge (Treasure et al., 2015; Zipfel, Giel, Bulik, Hay, & Schmidt, 2015).

In RCTs, treatment outcomes are compared after the end of treatment on a group level, an approach, which largely ignores outcome variation between individuals (Collins & Sayer, 2000). However, there has been a greater emerging interest in the processes of change in Anorexia Nervosa, with researchers finding that factors relating to how patients process and use treatment between sessions is related to treatment outcomes (Hartmann et al., 2016; Zeeck et al., 2016). It may be further argued that focusing on the processes of change will help to determine when the majority of symptom change is taking place (Kazdin & Nock, 2003). Analyzing the content of therapeutic sessions prior to this symptom change may enable researchers to identify which components of therapy are most important in affecting longer-term symptom change.

One characteristic of the therapeutic process is sudden gains (SGs); defined as sudden reductions in symptoms between two consecutive treatment sessions. SGs were first introduced by Tang and DeRubeis (1999) in the context of Cognitive Behavioral Therapy (CBT) for depression, who argued that SGs should be large (a) in absolute terms, (b) relative to symptom severity before the gain, and (c) relative to symptom fluctuation preceding and following the gain. They defined SGs using scores on the Beck

Depression Inventory (BDI; Beck & Steer, 1987):

> "A sudden gain occurred between session *N* and session *N* + 1 if (a) the gain was at least 7 BDI points ($BDI_N - BDI_{N+1} \geq 7$); (b) the gain represented at least 25% of the pregain session's BDI score ($BDI_N - BDI_{N+1} \geq 0.25 \times BDI_N$), and (c) the mean BDI score of the three therapy sessions before the gain (sessions *N* – 2, *N* – 1, and *N*) was significantly higher than the mean BDI score of the three therapy sessions after the gain (sessions *N* + 1, *N* + 2, and *N* + 3) using a two-sample t test with alpha of .05." (Tang & DeRubeis, 1999)

More than 50% of treatment responders experienced SGs during treatment (Tang & DeRubeis, 1999) and those who experienced SGs were less depressed at post-treatment and 18-month follow-up compared to those who did not.

The original definition has been criticized and altered (Hardy et al., 2005). The 'absolute magnitude' criterion has been criticized for being arbitrary (Hofmann, Schulz, Meuret, Moscovitch, & Suvak, 2006); the 'relative magnitude' criterion has been criticized for assuming that symptom measures are ratio scales (Utzinger, Goldschmidt, Crosby, Peterson, & Wonderlich, 2016) and has been found to have minimal impact on SG selection (Tang, DeRubeis, Beberman, & Pham, 2005); and the 'stability' criterion has been criticized for precluding examination of gains between the first and second treatment sessions (Aderka, Nickerson, Bøe, & Hofmann, 2012). However, despite variations in definitions, SGs have been consistently associated with better treatment outcomes (Aderka et al., 2012).

The majority of studies on SGs have focused on anxiety and depression. Only one study has assessed the role of SGs in eating disorders, measured using the Change in Eating Disorder Symptoms Scale (CHEDS; Spangler, 2010), in cognitive behavioral therapy for bulimic eating disorders. SGs were defined using two criteria: the 'absolute

magnitude criterion required a decrease of at least 12 points on the CHEDS, and the 'stability criterion' was based on that originally used by Tang and DeRubeis (1999). Over 50% of participants had at least one SG, and those experiencing SGs had better CHEDS outcome scores post-treatment (Cavallini & Spangler, 2013). To date, no study has assessed whether SGs occur in patients with AN during treatment, and, if so, whether these are related to treatment outcomes.

There is also debate regarding the importance of the timing of SGs. Tang and DeRubeis (1999) found that SGs typically occurred early, around session 5, and there is some evidence to suggest that SGs early in treatment are more relevant to treatment outcomes (Lutz, Bachmann, Tschitsaz, Smart, & Lambert, 2007; Stiles et al., 2003). However, there is considerable variation in the definition of what constitutes 'early'. Furthermore, in a large meta-analysis, Aderka et al. (2012) found no significant differences in outcomes between those who experienced SGs early vs. later in treatment.

Although SGs have been associated with better outcomes post-treatment, this is not always maintained at longer-term follow-up (Clerkin, Teachman, & Smith-Janik, 2009). In addition, SGs are often only associated with improvements in the primary outcome measure, which is typically the same measure used to define SGs (Aderka et al., 2012). This may be due to definitions of SGs being too lenient, such that individuals who experience SGs are simply those who improve on the target measure during treatment. Additional criteria may therefore be required in order to ensure that SGs are truly sudden and are measuring more than simply symptom change during treatment.

There has been considerable debate regarding how SGs should be defined and measured in the context of AN (Utzinger et al., 2016). The current study therefore aimed to operationalize SGs, using data from a clinical trial of AN. We based our criteria on

those proposed by Tang and DeRubeis (1999), adding an additional criterion of 'suddenness'. We aimed to explore whether there are SGs in BMI during treatment of AN; describe the characteristics of SGs, including their frequency, magnitude and timing; and determine whether the proportion and timing of SGs predicts outcomes at 6, 12 and 24 months follow-up. We also aimed to determine whether baseline variables (including demographic and clinical characteristics) differed between those who did and did not experience SGs during treatment.

METHODS

Data Source

The current study used data from The Maudsley Outpatient Study of Treatments for Anorexia Nervosa and Related Conditions (MOSAIC; Schmidt et al., 2015), and this trial's 2-year follow-up (Schmidt et al., 2016); a large multicenter, two-arm superiority RCT of adult outpatients with broadly defined AN. Patients were randomized to receive one of two treatments: the Maudsley Model of Anorexia Nervosa Treatment for Adults (MANTRA), an empirically based cognitive-interpersonal manualized treatment targeting factors thought to maintain symptoms of AN (Schmidt & Treasure, 2006; Treasure & Schmidt, 2013); or Specialist Supportive Clinical Management (SSCM), a manualized treatment, which involves providing information, advice, and encouragement in a supportive therapeutic manner (McIntosh et al., 2006). All participants were offered at least 20 individual therapy sessions (patients with a BMI ≤ 15kg/m^2 were offered 30 individual sessions), 4 monthly follow-up sessions, and 2 additional sessions with a close other. Ethical approval for the trial was obtained from the Central London Research Ethics Committee (REC) 4, Royal Free Hospital, London, National Health Service; REC Reference: 10/H0714/9. Full details of this trial are

available elsewhere (Schmidt et al., 2013, 2015, 2016).

Participants

MOSAIC participants were recruited from four outpatient ED services in the United Kingdom. Information regarding inclusion/exclusion criteria and participant flow through the MOSAIC trial is available elsewhere (Schmidt et al., 2015, 2016). Exclusion criteria for this study were having fewer than 4 weight measurements and no height measurement. Information regarding availability of weight data is presented in figure 1. As can be seen from figure 1, almost 30% of MOSAIC patient files were missing. This is due to factors such as patients failing to start treatment (n=10), files being lost, and therapist non-compliance with record keeping, with two therapists accounting for 10 missing files. Comparisons of demographic and clinical characteristics of missing MOSAIC patients and the present subsample are presented in table 1, indicating minimal differences between these two subsamples.

Outcome Measures

The primary outcome in the MOSAIC trial was BMI (kg/m^2). Other secondary outcome measures included: ED psychopathology, measured using the Eating Disorders Examination (EDE) Interview (Fairburn, Cooper, & O'Connor, 2008), or the questionnaire form of this assessment (EDE-Q); general psychopathology, measured using the Depression, Anxiety and Stress Scale-21(DASS-21; (Lovibond & Lovibond, 1995)); obsessive-compulsive symptomatology, measured using the Obsessive Compulsive Inventory-Revised (OCI-R; (Foa et al., 2002)); and psychosocial impairment, measured using the Clinical Impairment Assessment (CIA; Bohn & Fairburn, 2008). Information regarding reliability and validity of these measures can be found elsewhere

(see Schmidt et al., 2015). EDE and EDE-Q data in the MOSAIC trial were combined into a single EDE outcome measure. BMI and EDE was measured at baseline and 6, 12 ad 24 months follow-up. All other outcomes were measured at baseline and 24 months follow-up.

Procedure

Session-by-session weight measures were extracted from patients' files. Weight measures taken more than 365 days after baseline (n=84) were excluded to ensure that SGs were not measured after the 12-month follow-up. The researchers recorded the session at which each weight measure was taken and the number of days between baseline and each session. All other data (demographic characteristics, clinical details, and outcome measures described above) were extracted from MOSAIC trial databases.

Defining Sudden Gains

Weight measures were transformed into BMI (weight/height)^2. SGs were measured in terms of change in BMI between one treatment session (pre-gain session, N) and the following session (post-gain session, N+1). Meeting the criterion of absolute magnitude required an average between-session increase of ≥0.183kg/m^2 per week. This criterion is based on the NICE guidelines' recommendation of an average weight gain of 0.5kg /week for adults with anorexia nervosa treated in outpatient settings (National Collaborating Centre for Mental Health, 2004), and was calculated using the mean height of our sample. Although it may be argued that a weight gain expected to be the norm within the NICE guidelines is not 'large', the mean weight gain during treatment for AN is much lower than this recommended amount (Hartmann, Weber, Herpertz, Zeeck & German Treatment Guideline Group for Anorexia N., 2011). Expecting a weight

gain of more than this amount may therefore be considered unrealistic. The second criterion of relative magnitude was dropped, as it has been shown to have minimal impact on SG selection (Tang et al., 2005). Meeting the third criterion of stability required the mean BMI of the three measures taken during sessions N-2, N-1 and N to be significantly lower than the mean BMI of sessions N+1, N+2 and N+3, as calculated using an independent samples t-test ($\alpha = 0.05$). To maximize data inclusion for SG calculation, when BMI gains occurred between sessions where there were two pre-gain and/or two post-gain measures, t-tests were calculated using these values. Finally, we added an additional criterion of 'suddenness', which required the rate of BMI increase between sessions N and N+1 to be ≥1.5 times the rate of BMI change between sessions N-1 and N.

Due to considerable variation in the number of weight measures per participant, the proportion of SGs was calculated by dividing patients' total number of SGs by that patients' maximum possible number of SGs (i.e. total number of weight measure – 3).

Statistical Analysis

To identify baseline characteristics associated with SGs, participants were split into those who did and did not experience SGs. Independent samples t-tests or chi-squared tests were used to compare demographic characteristics and baseline clinical details of the two groups. Nonparametric tests were used where appropriate.

Multiple linear regression analyses were used to investigate whether the proportion of SGs predicted better long-term outcomes. Outcome measures included the difference in BMI and EDE scores between baseline and 6, 12 and 24 months post-randomization, and the difference in DASS-21, OCI-R, and CIA scores between baseline and 24 months post-randomization. In order to maximize statistical power, data were

combined across treatment arms. In doing this, we are assuming that any associations between the proportion and timing of SGs and long-term treatment outcomes are not dependent on the type of therapy. Baseline measures of bingeing, vomiting and laxative use were included as confounders due to the effects of these variables on weight, for example due to water retention (Rigaud, Boulier, Tallonneau, Brindisi, & Rozen, 2010). Finally, although not found to have significant effects on the outcome in the primary trial analyses (Schmidt et al., 2015, 2016) treatment effects were also allowed for by adding treatment group into the regression model.

To determine whether the timing of the first SG predicts treatment outcomes, regression analyses were repeated for just those individuals experiencing SGs, with the independent variable: number of days between randomization and the first SG.

RESULTS

Describing Sudden Gains

A total of 1697 treatment sessions were analyzed for the identification of SGs across all participants, creating 1607 between-session comparisons. Within our final sample, the median number of weight measurements was 20 (range = 4 to 35), and the median length of time between measures was 7 days (range = 3 to 140).

There were 327 between-session comparisons meeting the 'absolute magnitude' criterion, 122 of which also met the 'stability' criterion. Adding the criterion of 'suddenness' left 102 SGs amongst 55 participants. Thus, 61.8% of all participants experienced at least 1 SG in BMI during therapy.

We examined the length, frequency, magnitude and timing of SGs during both treatments. The median number of days between pregain and postgain weight measures was 7 (range = 5 to 49) for both treatments. The median number of SGs

experienced by patients in both treatment groups was 1 (MANTRA range = 0 to 8; SSCM range = 0 to 5). A graph showing the frequency distribution of the number of SGs can be found in the online supplement.

The median change in weight during periods defined as a SG was 0.81kg/week (range = 0.47 to 6.61): 0.83kg/week (range = 0.50 to 6.61) for MANTRA and 0.80kg/week (range = 0.47 to 1.52) for SSCM. This was equal to a BMI increase of 0.32 kg/m^2 per week for MANTRA (range = 0.19 to 2.43), and 0.31 kg/m^2 per week for SSCM (range = 0.19 to 0.77). The median length of time between baseline and a patient's first SG was 107 days (range = 29 to 259) for MANTRA, and 91.5 days (range = 27 to 292) for SSCM. The majority (73.5%) of SGs were experienced during the first 6 months of treatment, and 94.1% occurred within the fist 9 months.

SGs occurred most often after session 12 for MANTA and after sessions 5, 9, or 12 for SSCM (see figure 2). Three participants experienced a SG after their first session with a close other (session CO1).

Characteristics of Individuals Experiencing Sudden Gains

Those who had at least one SG had spent significantly more years in education (median=17.0, range=12.0-20.0) than those not experiencing SGs (median=15.0, range=8.0-20.0), $U = 453.00$, $p < .007$**, $r = 0.31$. Those who had at least one SG had a significantly lower DASS-21 score at baseline (mean=27.4, SD=10.6) compared to those not experiencing SGs (mean=32.9, SD=11.3), $t(85) = 2.27$, $p < .026$*, $d = 0.50$., 95% CI [0.67, 10.19]. Comparing demographic and clinical characteristics of those who did and did not experience SGs revealed no other significant differences (see table 2, online supplement).

Summary Statistics of Outcome Measures

Table 3 shows the mean BMI and EDE global scores, as measured at baseline and 6, 12 and 24 months post-randomization and all other outcomes measured at baseline and 24 months post-randomization, separately for individuals who did and did not experience SGs.

Associations Between The Proportion of Sudden Gains and Long-Term Treatment Outcomes

The proportion of SGs was a significant positive predictor of BMI change between baseline (0 months) and 6, 12 and 24 months post-randomization, with large sizes of effect (table 4). Our model did not explain a significant amount of the variance in the change in any other outcome measures.

Associations Between The Number of Days to a Patient's First Sudden Gain and Long-Term Treatment Outcomes

The number of days between randomization and a patient's first SG was a significant negative predictor of BMI change between baseline (0 months) and 6 and 12 months post-randomization, with moderate sizes of effect. However, the number of days between randomization and a patient's first SG was not a significant predictor of BMI change between baseline and 24 months post-randomization (table 5). The number of days between randomization and a patient's first SG was not a significant predictor of change in any of the MOSAIC trial secondary outcomes. It should also be noted that there was not a significant correlation between the proportion of SGs and the number of days to a patient's first SG ($p<0.13$).

DISCUSSION

This study was the first to define and operationalize SGs in the context of AN. We found that 61.8% of patients experienced at least one SG in BMI during treatment. As hypothesized, a larger proportion of SGs predicted larger increases in BMI between baseline and 6, 12 and 24 months post-randomization. Amongst those experiencing at least one SG during treatment, fewer days between randomization and a patient's first SG predicted a larger increase in BMI between baseline and both 6 and 12 months post-randomization.

Despite including an additional criterion of suddenness, the percentage of patients experiencing SGs in the current study is somewhat larger than the 40-50% that has previously been reported during psychological treatment for anxiety and depression (Aderka et al., 2012). This may be due to differences in outcome measures used to define SGs. Alternatively; it may be that the criterion of absolute magnitude used in this study is more lenient than previous definitions. The present study required an average between-session gain of 0.183kg/m^2 per week. It could be argued that this is not a 'large' increase in weight, and therefore does not adhere to the criterion of 'large absolute magnitude'. However, substantial weight gain does not typically occur over short periods of time, and rapid weight gain can be dangerous, especially in the context of AN (Utzinger et al., 2016). We therefore adapted the SG criteria in order to use BMI as the outcome variable, ensuring that we were adhering to NICE guidelines and not requiring a weight increase that was dangerous or unrealistic.

Compared to those experiencing no SGs, individuals who experienced at least one SG had spent significantly longer in education, and had a significantly lower DASS-21 total score at baseline. The reasons for the association with education are unclear, however lower depression scores have previously been associated with more favorable

treatment outcomes in ED treatment (Wild et al., 2016; Vall & Wade, 2015). This study therefore supports these findings, suggesting that lower depressions scores are associated with better treatment outcomes during the outpatient treatment of AN. There were no other significant differences between those who did and did not experience SGs during treatment, suggesting that other demographic and clinical characteristics cannot be used to reliably differentiate between individuals who do and do not experience SGs.

As hypothesized, a larger proportion of SGs predicted larger increases in BMI between baseline and 6, 12 and 24 months post-randomization. Importantly, there were no differences in BMI at baseline between those who did and did not experience SGs, suggesting that those experiencing SGs did not have a lower BMI at baseline and then regress towards the mean, as has been previously suggested (Konig, Karl, Rosner, & Butollo, 2014). However, it may not be surprising that individuals who experienced more SGs in BMI during treatment had larger increases in BMI between baseline and follow-up, as SGs and outcomes were both measured from baseline, and therefore covered the same time period. Nevertheless, our results are clinically useful, as they provide insight into the process of change during the treatment of AN, and suggest that change in BMI during AN treatment often occurs suddenly between treatment sessions. Furthermore, our findings suggest that the SGs in BMI that are experienced during treatment are maintained at follow-up and are therefore preferable during treatment.

SGs most often occurred during the middle phases of treatment, a time when therapists typically work on eliciting change. This was especially true for MANTRA patients, suggesting that there may be something specific about these middle sessions, which elicit SGs. The focus on clinical formulation and the creation of treatment plans may be responsible for these SGs in BMI. These findings regarding the timings of SGs

contradict some previous suggestions that SGs most often occur early during treatment. Within the present study, when defining SGs, there was no distinction made between those that occurred early and later during treatment, due to the considerable variation in the timing of SGs and the controversy regarding the definition of 'early'. This variation in timing of SGs within the present sample may be due to the fact that, although treatments were manualized, there was considerable flexibility regarding the specific content of each therapy session. Thus, it is possible that certain topics covered during treatment coincided with SGs, but were covered in different sessions for different patients. Alternatively, it is possible that factors such as motivation to change, or therapeutic alliance were associated with SGs, rather than the content of treatment sessions.

In line with previous findings (Lutz et al., 2007), amongst those experiencing at least one SG during treatment, fewer days between randomization and a patient's first SG predicted larger increases in BMI between baseline and both 6 and 12 months post-randomization. It may consequently be suggested that individuals who experience SGs earlier during treatment have better treatment outcomes. This association does not appear to be mediated by the total number of SGs experienced during treatment, as there was no significant correlation between proportion of SGs and the number of days between randomization and patients' first SG. Our findings did not support the hypothesis that the number of days between randomization and a patient's first SG would be predictive of BMI change between baseline and 24 months post-randomization.

Limitations and Conclusions

This study has several limitations. Firstly, almost 30% of MOSAIC patient files were

missing. This was partly due to patients dropping out of the study prior to the start of treatment and therapist non-compliance with record keeping. However, it is also possible that files may have been missing for reasons, such as patients not attending therapy sessions. Similarly, there was a large variation in the number of weight measures recorded for each participant, and lack of consistency in the timing of these recordings, making it difficult to precisely investigate the trajectory of weight change. This complicated the calculations of SGs, requiring us to use the proportion of SGs rather than the raw number of SGs, making interpretation of these findings more challenging.

The MOSAIC study was not designed to investigate SGs and the data was therefore not optimal for the current study. Firstly, weight was the only variable to be measured at each session, limiting our choice as to how to define SGs. In addition, the number of days between weight measures varied considerably. Large gaps between measurements make it impossible to determine the true suddenness of SGs. Other possible reasons for missing weight data include the fact that there was a large variability in the number of treatment sessions attended by patients during the MOSAIC trial (Schmidt et al., 2015); and patients sometimes refused to be weighed. It is likely that this is not a random subsample of patients, further highlighting the challenges of missing data. We encourage future researchers to more rigorously collect session-by-session data for multiple variables in order to investigate SGs in BMI and other, psychological symptoms of AN.

Despite these limitations, the present study progresses research into the treatment of AN by suggesting that there are SGs in BMI during treatment, and that these are related to better long-term outcomes. Further research is, however, required to determine which components of therapy most often precede SGs in order to identify

aspects of therapy that are most crucial in affecting change. Future research could also explore whether there are SGs in psychological symptoms of AN, and, if present, whether these are related to better outcomes. This would also allow comparisons to be made with research into SGs in CHEDS scores during treatment for bulimia nervosa (Cavallini & Spangler, 2013).

In sum, the present study was the first to investigate SGs during the treatment of AN. Our findings provide support for previous research in other psychiatric disorders suggesting that SGs are associated with better treatment outcomes (Aderka et al., 2012) and suggest that patients show SGs in weight during treatment. Future research investigating factors that precede these SGs may help researchers to identify the components of therapy that are most important in affecting change in AN, and thereby contributing to the development of more effective treatments.

REFERENCES

Aderka, I. M., Nickerson, A., Bøe, H. J., & Hofmann, S. G. (2012). Sudden gains during psychological treatments of anxiety and depression: a meta-analysis. *Journal of Consulting and Clinical Psychology*, *80*(1), 93–101. doi: 10.1037/a0026455

Beck, A. T., & Steer, R. A. (1987). *Manual for the Beck Depression Inventory.* San Antonio, TX: The Psychological Corporation.

Bohn, K., & Fairburn, C. G. (2008). The Clinical Impairment Assessment Questionnaire (CIA 3.0). In C. G. Fairburn (Ed.), *Cognitive behavioral therapy for eating disorders* (pp. 315–318). New York: NY: Guilford Press.

Cavallini, A. Q., & Spangler, D. L. (2013). Sudden gains in cognitive-behavioral therapy for eating disorders. *International Journal of Cognitive Therapy*, *6*, 292–310. doi: 10.1521/ijct.2013.6.3.292

Clerkin, E. M., Teachman, B. A., & Smith-Janik, S. B. (2009). Sudden gains in group cognitive-behavioral therapy for panic disorder. *Behaviour Research and Therapy*, *46*(11), 1244–1250. doi: 10.1016/j.brat.2008.08.002

Collins, L. M., & Sayer, A. G. (2000). Growth and change in social psychology research: Design, measurement, and analysis. In H. Reis & C. Judd (Eds.), *Handbook of Research in Social Psychology* (pp. 478–495). Cambridge: Cambridge University Press (CUP).

Fairburn, C. G., Cooper, Z., & O'Connor, M. (2008). Eating disorder examination. In C. G. Fairburn (Ed.), *Cognitive behavior therapy and eating disorders* (pp. 270–308). New York: NY: Guilford Press.

Foa, E. B., Huppert, J. D., Leiberg, S., Langner, R., Kichic, R., Hajcak, G., & Salkovskis, P. M. (2002). The Obsessive-Compulsive Inventory: Development and validation of a short version. *Psychological Assessment*, *14 (4)*, 485–496. doi: 10.1037/1040-3590.14.4.485

Gatt, L., Jan, S., Mondraty, N., Horsfield, S., Hart, S., Russell, J., …Essue, B. (2014). The household economic burden of eating disorders and adherence to treatment in Australia. *BMC Psychiatry*, *14*(1), 338. doi: 10.1186/s12888-014-0338-0

Hardy, G. E., Cahill, J., Stiles, W. B., Ispan, C., Macaskill, N., & Barkham, M. (2005). Sudden gains in cognitive therapy for depression: a replication and extension. *Journal of Consulting and Clinical Psychology*, *73*(1), 59–67. doi: 10.1037/0022-006X.73.1.59

Hartmann, A., Weber, S., Herpertz, S., Zeeck, A., German Treatment Guideline Group for Anorexia Nervosa (2011). Psychological treatment for anorexia nervosa: a meta-analysis of standardized mean change. *Psychotherapy and Psychosomatics, 80*(4), 216-226. doi: 10.1159/000322360

Hartmann, A., Zeeck, A., Herzog, W., Wild, B., De Zwaan, M., Herpertz, S., …Zipfel, S. (2016). The intersession process in psychotherapy for anorexia nervosa: Characteristics and relation to outcome. Journal of Clinical Psychology, *72*, 861–879. doi:10.1002/jclp.22293

Hofmann, S. G., Schulz, S. M., Meuret, A. E., Moscovitch, D. A., & Suvak, M. (2006). Sudden gains during therapy of social phobia. *Journal of Consulting and Clinical Psychology*, *74*(4), 687–697. doi: 10.1038/jid.2014.371

Kazdin, A. E., & Nock, M. K. (2003). Delineating mechanisms of change in child and

adolescent therapy: methodological issues and research recommendations. *Journal of Child Psychology and Psychiatry*, *44*, 1116–1129. doi: 10.1111/1469-7610.00195

Konig, J., Karl, R., Rosner, R., & Butollo, W. (2014). Sudden gains in two psychotherapies for posttraumatic stress disorder. *Behaviour Research and Therapy*, *60*, 15–22. doi: 10.1016/j.brat.2014.06.005

Lovibond, S. H., & Lovibond, P. F. (1995). *Manual for the Depression Anxiety Stress Scales* (2nd ed.). Sydney, Australia: Psychology Foundation.

Lutz, W., Bachmann, F., Tschitsaz, A., Smart, D. W., & Lambert, M. (2007). Temporal and sequential patterns of nonlinear change psychotherapy. The phenomenon of sudden gains . *Zeitschrift Für Klinische Psychologie Und Psychotherapie*, *36*, 261–269.

McIntosh, V. V., Jordan, J., Luty, S. E., Carter, F. A., McKenzie, J. M., Bulik, C. M., & Joyce, P. R. (20006). Specialist supportive clinical management for anorexia nervosa. *International Journal of Eating Disorders*, *39*, 625–632. doi: 10.1002/eat.20297

National Collaborating Centre for Mental Health. (2004). Eating disorders: core interventions in the treatment and management of anorexia nervosa, bulimia nervosa and related eating disorders. In *NICE Clinical Guideline*.

Rigaud, D., Boulier, A., Tallonneau, I., Brindisi, M. C., & Rozen, R. (2010). Body fluid retention and body weight change in anorexia nervosa patients during refeeding. *Clinical Nutrition*, *29*(6), 749–755. doi: 10.1016/j.clnu.2010.05.007

Schmidt, U., Magill, N., Renwick, B., Keyes, A., Kenyon, M., DeJong, H., … Landau, S. (2015). The Maudsley Outpatient Study of Treatments for Anorexia Nervosa and

Related Conditions (MOSAIC): Comparison of the Maudsley Model of Anorexia Nervosa Treatment for Adults (MANTRA) with Specialist Supportive Clinical Management (SSCM) in outpatients with broadly defined anorexia nervosa: A randomized controlled trial. *Journal of Consulting and Clinical Psychology*, *83*(4), 796–807. doi: 10.1037/ccp0000019

Schmidt, U., Renwick, B., Lose, A., Kenyon, M., Dejong, H., Broadbent, H., … Landau, S. (2013). The MOSAIC study - comparison of the Maudsley Model of Treatment for Adults with Anorexia Nervosa (MANTRA) with Specialist Supportive Clinical Management (SSCM) in outpatients with anorexia nervosa or eating disorder not otherwise specified, anorexia nervosa type: study protocol for a randomized controlled trial. *Trials*, *14*(1), 160. doi: 10.1186/1745-6215-14-160

Schmidt, U., Ryan, E. G., Bartholdy, S., Renwick, B., Keyes, A., O'Hara, C., … Treasure, J. (2016). Two-year follow-up of the MOSAIC trial: A multicenter randomised controlled trial comparing two psychological treatments in adult outpatients with broadly defined anorexia nervosa. *International Journal of Eating Disorders*, *49*(8), 793–800. doi: 10.1002/eat.22523

Schmidt, U., & Treasure, J. (2006). Anorexia nervosa: valued and visible. A cognitive-interpersonal maintenance model and its implications for research and practice. *British Journal of Clinical Psychology*, *45*, 343–366. doi: 10.1348/014466505X53902

Spangler, D. L. (2010). The change in eating disorder symptoms scale: Scale development and psychometric properties. *Eating Behaviors*, *11*, 131–137. doi: 10.1016/j.eatbeh.2009.12.003

Stiles, W. B., Leach, C., Barkham, M., Lucock, M., Iveson, S., Shapiro, D. A., …Hardy, G. E. (2003). Early sudden gains in psychotherapy under routine clinic conditions: practice-based evidence. *Journal of Consulting and Clinical Psychology*, *71*(1), 14–21. doi: 10.1037/0022-006X.71.1.14

Stuhldreher, N., Konnopka, A., Wild, B., Herzog, W., Zipfel, S., Löwe, B., & König, H. H. (2012). Cost-of-illness studies and cost-effectiveness analyses in eating disorders: A systematic review. *International Journal of Eating Disorders*, *45*(4), 476–491. doi: 10.1002/eat.20977

Tang, T., & DeRubeis, R. (1999). Sudden gains and critical sessions in cognitive-behavioral therapy for depression. *Journal of Consulting and Clinical Psychology*, *67*(6), 894–904. doi: 10.1037/0022-006X.75.3.404

Tang, T. Z., DeRubeis, R. J., Beberman, R., & Pham, T. (2005). Cognitive changes, critical sessions, and sudden gains in cognitive-behavioral therapy for depression. *Journal of Consulting and Clinical Psychology*, *73*(1), 168–172. doi: 10.1037/0022-006X.73.1.168

Treasure, J., & Schmidt, U. (2013). The cognitive-interpersonal maintenance model of anorexia nervosa revisited: a summary of the evidence for cognitive, socio-emotional and interpersonal predisposing and perpetuating factors. *Journal of Eating Disorders*, *1*(13), 1–10. doi: 10.1186/2050-2974-1-13

Treasure, J., Zipfel, S., Micali, N., Wade, T., Stice, E., Claudino, A., …Wentz, E. (2015). Anorexia nervosa. *Nature Reviews Disease Primers*, *1*(November), 1–21. doi: 10.1038/nrdp.2015.74

Utzinger, L. M., Goldschmidt, A. B., Crosby, R. D., Peterson, C. B., & Wonderlich, S. A. (2016). Are sudden gains important in the treatment of eating disorders? *International Journal of Eating Disorders*, *49*(1): 32-35. doi: 10.1002/eat.22458

Vall, E., & Wade, T. D. (2015). Predictors of treatment outcome in individuals with eating disorders: A systematic review and meta-analysis. *International Journal of Eating Disorders*, *48*(7), 946–971. doi: 10.1002/eat.22411

Wild, B., Friederich, H.-C., Zipfel, S., Resmark, G., Giel, K., Teufel, M., … Herzog, W. (2016). Predictors of outcomes in outpatients with anorexia nervosa – Results from the ANTOP study. *Psychiatry Research*, *244*, 45–50. doi: 10.1016/j.psychres.2016.07.002

Zeeck, A., Hartmann, A., Wild, B., De Zwaan, M., Herpertz, S., Burgmer, M., …Antop Study Group (2016). How do patients with anorexia nervosa 'process' psychotherapy between sessions? A comparison of cognitive-behavioral and psychodynamic interventions. *Psychotherapy Research, Nov 3,* 1-14. doi: 10.1080/10503307.2016.1252866

Zipfel, S., Giel, K. E., Bulik, C. M., Hay, P., & Schmidt, U. (2015). Anorexia nervosa: Aetiology, assessment, and treatment. *The Lancet Psychiatry*, *2*(12), 1099–1111. doi: 10.1016/S2215-0366(15)00356-9

FIGURE CAPTIONS

Figure 1. Diagram illustrating the availability of weight data at various assessment time points.

Figure 2. Frequency distribution of the session number preceding all sudden gains. CO1 = first session with a close other.

Online Supplement Figure. Frequency distribution of the number of sudden gains per patient.

Table 1.
Baseline Characteristics of the Missing MOSAIC sample and the present sample

	Missing Sample		Present Sample		Test
	N		*N*		
Demographic details					
Age, years (median [range])	53	25 (18-49)	89	23 (18-52)	$U = 2219.50$, $p = .56$
Men:Women (n)	53	1:52	89	2:87	FET, $p = 1.00$ (2-sided)
Years in education (median [range])	47	16 (12-20)	79	16 (8-20)	$U = 1787.00$, $p = .88$
In relationship (n [%])	50	15 (30.0)	88	35 (39.8)	$\chi^2_{(1)} = 1.32$, $p = .25$
Clinical details					
Treatment group (SSCM: MANTRA)	53	32:21	89	38:51	$\chi^2_{(1)} = 4.16$, $p = .04$*
Diagnosis (n [%])	53		89		$\chi^2_{(2)} = 2.11$, $p = .35$
AN-R		28 (52.8)		37 (41.6)	
AN-BP		15 (28.3)		27 (30.3)	
EDNOS		10 (18.9)		25 (28.1)	
BMI, kg/m^2 (mean [SD])	47	16.6 (1.2)	87	16.7 (1.3)	$t(132) = 0.42$, $p = .67$
Weight, kg (mean [SD])	48	44.6 (4.5)	88	45.6 (4.9)	$t(134) = -1.16$, $p = .25$
Binge (median [range])	50	0 (0-20)	83	0 (0-40)	$U = 2067.00$, $p = .85$
Vomit (median [range])	51	0 (0-20)	86	0 (0-40)	$U = 2138.50$, $p = .78$
Laxative use (median [range])	50	0 (0-7)	86	0 (0-21)	$U = 2002.00$, $p = .47$
Age at onset, years (median [range])	48	16.0 (5.0-43.0)	86	16.5 (2.0-44.0)	$U = 1856.50$, $p = .33$
Illness duration, years (median [range])	48	7.5 (1.0-37.0)	86	5.5 (0.5-34.0)	$U = 1850.00$, $p = .32$
Previous ED treatment (n [%])	52	29 (55.8)	89	51 (57.3)	$\chi^2_{(1)} = 0.06$, $p = .80$
EDE (mean [SD])	53	3.4 (1.4)	89	3.3 (1.2)	$t(140) = 0.47$, $p = .64$
DASS-21 (mean [SD])	51	32.2 (15.0)	87	29.5 (11.1)	$t(82.23) = 1.13$, $p = .26$
OCI-R (mean [SD])	51	25.6 (15.9)	88	22.4 (12.2)	$t(84.54) = 1.26$, $p = .21$
CIA (mean [SD])	51	33.8 (10.2)	89	31.8 (8.0)	$t(139) = 1.23$, $p = .21$
Current antidepressant medication (n [%])	52	55 (26.9)	88	41 (46.6)	$\chi^2_{(1)} = 5.31$, $p = .02$*

Note. *$p < 0.05$. The Missing Sample column includes participants who were in the MOSAIC sample, who were not included within the present sample. The Test column shows results of t-tests, Mann-Whitney U tests, Chi-Square tests, and Fisher's Exact Tests used to test for differences between the two groups. SSCM Specialist Supportive Clinical Management; MANTRA Maudsley Model of Anorexia Nervosa Treatment for Adults; AN-R anorexia nervosa, restricting type; AN-BP anorexia nervosa, binge eating/purging type; EDNOS Eating Disorder Not Otherwise Specified; BMI body mass index; EDE Eating Disorder Examination; DASS-21 Depression Anxiety Stress Scale; OCI-R Obsessive Compulsive Inventory–Revised; CIA Clinical Impairment Assessment.

Table 2.

Baseline Characteristics of participants with and without sudden gains

	SG		No SG		Test
	N		*N*		
Demographic details					
Age, years (median [range])	55	24.0 (18.0-45.0)	34	23.0 (18.0-52.0)	$U = 915.00$, $p = .87$
Men:Women (*n*)	55	0:55	34	2:32	FET, $p = .14$ (2-sided)
Years in education (median [range])	51	17.0 (12.0-20.0)	28	15.0 (8.0-20.0)	$U = 453.00$, $p = .007$**, $r = 0.31$
In relationship (*n* [%])	55	19.0 (34.5)	34	16.0 (47.1)	$\chi^2_{(1)} = 1.23$, $p = .27$
Clinical details					
Treatment group (SSCM:MANTRA)	55	20:35	34	18:16	$\chi^2_{(1)} = 2.36$, $p = .12$
Diagnosis (*n* [%])	55		34		$\chi^2_{(2)} = 0.13$, $p = .94$
AN-R		23 (41.8)		14 (41.2)	
AN-BP		16 (29.1)		11 (32.4)	
EDNOS		16 (29.1)		9 (26.5)	
BMI, kg/m² (mean [*SD*])	55	16.7 (1.3)	34	16.6 (0.9)	$t(87) = -0.33$, $p = .75$
Weight, kg (mean [*SD*])	54	45.6 (5.2)	34	45.7 (4.4)	$t(86) = 0.14$, $p = .89$
Binge (median [range])	55	0.0 (0.0-40.0)	34	0.0 (0.0-28.0)	$U = 906.00$, $p = .76$
Vomit (median [range])	55	0.0 (0.0-40.0)	34	0.0 (0.0-28.0	$U = 827.50$, $p = .29$
Laxative use (median [range])	55	0.0 (0.0-21.0)	34	0.0 (0.0-7.0)	$U = 893.00$, $p = .54$
Age at onset (median [range])	52	17.0 (9.0-44.0)	33	16.0 (11.0-34.0)	$U = 849.00$, $p = .82$
Illness duration, years (median [range])	53	5.0 (0.5-24.0)	33	6.0 (1.0-34.0)	$U = 807.50$, $p = .55$
Previous ED treatment (*n* [%])	55	31 (56.4)	34	20 (58.8)	$\chi^2_{(1)} = 0.15$, $p = .70$
EDE Global Score (mean [*SD*])	55	3.3 (1.2)	34	3.2 (1.3)	$t(87) = -0.31$, $p = .76$
DASS-21 Total Score (mean [*SD*])	54	27.4 (10.6)	33	32.9 (11.3)	$t(85) = 2.27$, $p = .026$*, $d = 0.50$
OCI-R Total Score (mean [*SD*])	55	21.7 (12.0)	33	23.5 (11.7)	$t(86) = 0.66$, $p = .51$
CIA Total Score (mean [*SD*])	55	31.7 (8.5)	34	32.1 (7.4)	$t(87) = 0.24$, $p = .81$
On antidepressant medication (*n* [%])	55	24 (43.6)	33	17 (51.5)	$\chi^2_{(1)} = 0.52$, $p = .48$

Note. *$p < 0.05$, **$p < 0.01$. The SG column includes participants experiencing ≥ 1 sudden gains (SGs). The No SG column includes participants who did not have any SGs. The Test column shows results of t-tests, Mann-Whitney U tests, Chi-Square tests, and Fisher's Exact Tests used to test for differences between the two groups. SSCM Specialist Supportive Clinical Management; MANTRA Maudsley Model of Anorexia Nervosa Treatment for Adults; AN-R anorexia nervosa, restricting type; AN-BP anorexia nervosa, binge eating/purging type; EDNOS Eating Disorder Not Otherwise Specified; BMI body mass index; EDE Eating Disorder Examination; DASS-21 Depression Anxiety Stress Scale; OCI-R Obsessive Compulsive Inventory–Revised; CIA Clinical Impairment Assessment; FET Fisher's Exact Test.

Table 3.
Summary statistics for all outcome measures presented separately for those who did and did not experience sudden gains during treatment

Measure	SG		No SG	
	N	Mean (SD)	*N*	Mean (SD)
BMI				
Baseline	55	16.73 (1.34)	34	16.65 (0.87)
6-months	51	17.81 (1.51)	29	16.44 (1.44)
12-months	49	18.72 (2.09)	27	16.83 (1.63)
24-months	48	18.97 (2.39)	21	17.00 (2.03)
EDE Global				
Baseline	55	3.23 (1.25)	34	3.32 (1.25)
6-months	52	2.45 (1.38)	29	2.54 (1.36)
12-months	49	2.16 (1.50)	26	2.44 (1.39)
24-months	48	2.09 (1.55)	21	2.89 (1.47)
EDE Restraint				
Baseline	55	3.63 (1.49)	34	3.75 (1.58)
24-months	48	1.58 (1.40)	21	2.39 (1.43)
EDE Shape Concern				
Baseline	55	3.34 (1.72)	34	3.67 (1.59)
24-months	48	2.28 (1.77)	21	2.83 (1.82)
EDE Weight Concern				
Baseline	55	3.14 (1.61)	34	3.20 (1.67)
24-months	48	2.15 (1.52)	21	2.89 (1.47)
EDE Eating Concern				
Baseline	55	2.80 (1.26)	34	2.67 (1.40)
24-months	48	2.58 (1.90)	21	3.24 (1.72)
DASS-21 Total				
Baseline	54	30.96 (10.38)	33	27.12 (11.94)
24-months	48	20.00 (11.80)	20	32.73 (13.27)
DASS-21 Depression				
Baseline	54	10.76 (4.88)	33	9.44 (5.00)
24-months	48	6.69 (5.51)	20	11.43 (6.13)
DASS-21 Anxiety				
Baseline	54	7.11 (4.04)	33	6.33 (4.66)
24-months	48	4.15 (3.54)	20	8.70 (4.97)
DASS-21 Stress				
Baseline	54	13.09 (4.09)	33	11.42 (4.13)
24-months	48	9.17 (4.98)	20	12.60 (4.19)
OCI-R Total				
Baseline	55	24.91 (11.87)	33	18.15 (11.79)
24-months	46	17.07 (13.47)	16	21.81 (13.24)
CIA Total				
Baseline	55	32.75 (7.98)	34	30.35 (8.06)
24 months	46	19.65 (11.71)	20	27.55 (11.64)

Note. Values in the SG column include only participants experiencing at least one sudden gain. Values in the No SG column include only participants who did not have any sudden gains throughout treatment.

Table 4

Results of multiple linear regression analyses exploring the associations between the proportion of sudden gains and long-term treatment outcomes

Outcome Measure	Regression Coefficient	Standard Error of Regression Coefficient	Standardised Regression Coefficient	*p*	*Adjusted* R^2
BMI					
0-6 months	11.80	2.59	0.49	<.001**	0.20**
0-12 months	16.25	3.35	0.52	<.001**	0.21**
0-24 months	18.72	4.04	0.54	<.001**	0.20**
EDE Global					
0-6 months	-0.27	3.14	-0.01	.94	0.03
0-12 months	1.21	3.58	0.04	.74	-0.03
0-24 months	-4.57	3.48	-0.17	.19	-0.02
EDE Restraint	-0.92	4.21	-0.03	.83	-0.06
EDE Shape Concern	1.23	4.88	0.03	.80	-0.02
EDE Weight Concern	-4.5	4.46	-0.13	.32	-0.03
EDE Eating Concern	-5.50	4.17	-0.17	.19	-0.01
DASS-21 Total	-44.87	35.51	-0.16	.21	0.06
DASS-21 Depression	-13.29	15.58	-0.11	.40	0.07
DASS-21 Anxiety	-20.68	11.93	-0.22	.09	0.08
DASS-21 Stress	-6.08	13.42	-0.06	.65	-0.04
OCI-R Total	-36.32	40.05	-0.13	.37	-0.06
CIA Total	-53.32	27.28	-0.26	.06	0.03

Note: ** $p < 0.01$. Unless otherwise stated, dependent variables were measured as the difference in symptoms between baseline and 24 months post-randomization.

Table 5.

Results of multiple linear regression analyses exploring the associations between the number of days between randomisation and a patient's first sudden gain and long-term treatment outcomes.

Outcome Measure	Regression Coefficient	Standard Error of Regression Coefficient	Standardised Regression Coefficient	*p*	*Adjusted* R^2
BMI					
0-6 months	-0.009	0.003	-0.38	.006**	0.19*
0-12 months	-0.012	0.004	-0.39	.01*	0.06
0-24 months	-0.009	0.005	-0.29	.07	-0.02
EDE Global					
0-6 months	0.007	0.004	0.24	.10	0.03
0-12 months	0.004	0.005	0.14	.36	-0.04
0-24 months	0.009	0.005	0.30	.06	0.03
EDE Restraint	0.009	0.005	0.28	.08	-0.01
EDE Shape Concern	0.010	0.006	0.24	.13	-0.01
EDE Weight Concern	0.010	0.005	0.29	.07	0.04
EDE Eating Concern	0.007	0.005	0.22	.16	-0.03
DASS-21 Total	0.031	0.038	0.13	.43	0.00
DASS-21 Depression	0.011	0.017	0.10	.52	0.11
DASS-21 Anxiety	-0.002	0.013	-0.03	.86	-0.10
DASS-21 Stress	0.022	0.015	0.23	.15	-0.04
OCI-R Total	0.067	0.044	0.24	.14	-0.04
CIA Total	0.021	0.032	0.10	.52	-0.04

Note: $^{**}p < 0.01$ $^{*}p < 0.05$. Unless otherwise stated, dependent variables were measured as the difference in symptoms between baseline and 24 months post-randomisation.